

J.J. Bénet¹M. Sanaa¹B. Dufour²B. Toma¹

Méthodologie des enquêtes en épidémiologie animale

BÉNET (J.J.), SANAA (M.), DUFOUR (B.), TOMA (B.). Méthodologie des enquêtes en épidémiologie animale. *Revue Elev. Méd. vét. Pays trop.*, 1993, 46 (3) : 403-422

Dans cet article, les auteurs passent en revue les bases méthodologiques des enquêtes, tant descriptives qu'analytiques, en épidémiologie animale. Au cours de cette synthèse, ils précisent les définitions des principaux termes utilisés en épidémiologie, ainsi que les modalités d'évaluation de la qualité de ces enquêtes.

Mots Clefs : Enquête - Méthode - Epidémiologie - Echantillonnage - Evaluation.

INTRODUCTION

En épidémiologie animale, comme en épidémiologie des maladies de l'Homme, des enquêtes sont utilisées pour décrire la situation d'une maladie (épidémiologie descriptive), ou pour tenter de déterminer les facteurs qui favorisent ou conditionnent son apparition (épidémiologie analytique).

Ces deux types d'enquêtes, répondant à des objectifs différents, ont des caractéristiques elles-mêmes différentes qu'il importe de bien distinguer, et doivent répondre à des critères de qualité adaptés aux objectifs.

Elles constituent de véritables modèles de référence à partir desquelles peuvent être déclinées des applications épidémiologiques extrêmement diverses. Leur spécificité réside soit dans l'objet d'étude, soit dans les buts visés, et non dans leurs principes méthodologiques.

Bien que différentes dans leurs caractéristiques, ces enquêtes partagent le même principe de quantification de phénomènes de santé dans des populations, par la mesure de la fréquence de ces phénomènes à l'aide d'indicateurs, et en collectant l'information nécessaire à cette mesure par des méthodes d'échantillonnage.

Dans les lignes qui suivent, seront étudiés les objectifs, les caractéristiques et les critères de qualité de ces deux types d'enquêtes, après avoir défini pour chacune d'elles, le secteur de l'épidémiologie dans lequel elle s'insère. On discutera pour finir de l'évaluation des enquêtes épidé-

mologiques. Au préalable, il faut présenter leur communauté méthodologique, tenant à l'élaboration d'un indicateur approprié, et à la collecte de l'information.

Bases des enquêtes épidémiologiques

Une enquête est une "recherche méthodique d'informations reposant notamment sur des mesures, des questions et des témoignages" (12). Son équivalent anglais est : "survey". Elle constitue une modalité particulière d'étude ("study"), qui, elle, consiste en la "recherche méthodique d'informations visant à connaître et/ou à comprendre" (12). Ce dernier terme est plus général, et ne préjuge pas des modalités de récolte de l'information. Il englobe le précédent. Il est fréquent en français d'utiliser le terme enquête, même si les modalités de réalisation ne correspondent pas à la définition, comme dans le cas de la valorisation de prélèvements disponibles dans un laboratoire. Il arrivera souvent aux auteurs d'utiliser le terme d'étude par la suite de l'exposé.

Les enquêtes épidémiologiques ne constituent que la mise en application des principes de base de la quantification d'un phénomène de santé dans une population. Il est indispensable d'en bien comprendre les étapes clés pour pouvoir ensuite aborder avec profit l'étude des enquêtes proprement dites.

Les indicateurs, outils de la quantification en épidémiologie

Pour quantifier un phénomène de santé dans une population, la méthode consiste à effectuer des dénombrements, en vue d'établir une fréquence, en effectuant le choix de l'incidence ou de la prévalence, et en les utilisant sous la forme d'indicateurs.

Etapas préalables à l'élaboration d'un indicateur épidémiologique

Les dénombrements

Ces dénombrements doivent être définis sous les deux aspects indispensables à leur réalisation :

1. Ecole nationale vétérinaire d'Alfort, Laboratoire d'épidémiologie et de gestion de la santé animale, 94704 Maisons-Alfort cedex, France.

2. Centre d'études vétérinaires et alimentaires, Direction générale, 94700 Maisons-Alfort Cedex, France.

- définition du "cas" : définition qualitative précise de ce que l'on veut dénombrer : cas clinique ; test expérimental (et son résultat) ; autre information ;

- définition de l'unité statistique : animal, cheptel, lot. Ce point a des répercussions essentielles sur la quantification du phénomène mesuré.

Déterminer une fréquence

On doit distinguer la fréquence absolue (nombre d'événements par intervalle de temps), et la fréquence relative (fréquence précédente divisée par le nombre d'occurrences possibles).

Détermination de l'effet du temps sur les dénombrements

L'intervention du temps dans les dénombrements conduit à distinguer deux types d'événements : les états à un moment donné ou pendant une période donnée (prévalence) et les changements d'état pendant la période considérée (incidence).

La prévalence est "le nombre total de cas ou de foyers d'une maladie, dans une population déterminée, au cours d'une période donnée ou à un instant donné" (12). Elle intègre à la fois l'apparition de nouveaux cas, et la durée de la maladie, puisque les cas apparus dans la ou les périodes précédentes et qui sont toujours à l'état de cas dans la période considérée sont comptabilisés avec les cas apparus pendant cette dernière période. Elle représente une fréquence globale de la maladie pendant une période déterminée, par exemple la prévalence annuelle. Elle ne renseigne pas sur la dynamique de la maladie, mais elle informe le responsable d'un programme de santé sur l'impact qu'elle a sur la population, et par conséquent sur l'ampleur des moyens à mettre éventuellement en oeuvre pour intervenir sur ces cas.

Par définition, l'incidence est "le nombre de cas ou de foyers nouveaux d'une maladie, dans une population déterminée, au cours d'une période donnée" (12). Elle permet de mesurer l'apparition d'une maladie chez des individus susceptibles (ou à risque) durant une période donnée. C'est elle qui permet de caractériser la forme épidémiologique d'une maladie. Elle joue un rôle déterminant dans les études explicatives.

D'un point de vue pratique, incidence et prévalence ne sont pas déterminées dans les mêmes conditions. Pour estimer la prévalence, on a besoin d'une seule détermination par animal, alors que pour l'incidence, il faut au moins deux mesures, une au début de la période d'observation (pour déterminer les animaux à risque) et d'autres pendant la même période.

L'incidence présente l'inconvénient d'être fortement influencée par les mouvements de sortie de la population (soustraction au risque), et par le fait que les animaux sont exposés au risque pendant des périodes de temps variables (addition d'animaux à risque). Ces difficultés

conduisent à introduire une dimension temporelle à l'incidence et à utiliser la notion d'animaux-temps (ou personne-temps en épidémiologie humaine)(voir plus loin).

Les indicateurs : des rapports

Définition

Un indicateur (de santé) est une "variable reflétant l'un des aspects de l'état de santé d'une population" (12).

Les étapes précédemment définies ne sont qu'un préliminaire à l'élaboration d'un indicateur (fig. 1). En effet, il faut éliminer les répercussions de différences d'effectifs de populations ; pour cela, les dénombrements sont traités sous forme de rapports, qui constituent finalement les indicateurs.

Le dénominateur détermine la nature de ce rapport : proportion, taux, ratio (figure 1), et par conséquent autant de types d'indicateurs différents. Il n'est pas exagéré de dire que tout le secret d'un indicateur réside dans le choix du bon dénominateur.

INDICATEURS	
Les étapes d'élaboration d'un indicateur	Différents indicateurs
1- Définition du cas à dénombrer	$\text{Proportion} = \frac{n^+}{(n^+ + n^-)}$ <p>(→ pourcentage, %)</p> $\text{Taux} = n^+ / \text{susceptibles soumis au risque et par unité de temps}$ $\text{Ratio} = \frac{a}{b}$
2- Choix de l'unité statistique	
3- Choix de l'intervalle de temps	
4- Incidence ou prévalence	
5- Choix du dénominateur	

Figure 1 : Les étapes de l'élaboration d'un indicateur épidémiologique, et les différents types d'indicateurs épidémiologiques. n^+ , n^- = individus présentant ou non (respectivement) le caractère étudié. a, b = effectifs appartenant à des ensembles distincts.

Proportion (figure 1)

Une proportion est un "rapport (ou quotient) entre deux grandeurs, dont le numérateur est inclus dans le dénominateur" (12) ; par exemple : la proportion de mâles dans une population.

Une proportion n'a pas de dimension, et ses valeurs sont comprises entre 0 et 1. Elle est souvent exprimée en pourcentage (c'est-à-dire multipliée par cent). C'est le cas de la prévalence. La proportion, le pourcentage sont donc des fréquences relatives.

Taux (figure 1)

Au sens large, il s'agit du "nombre d'événements observés par unité de temps, rapporté à la population soumise au risque" (12) ; par exemple : le taux de prévalence

annuelle ; le taux de prévalence instantanée au 31 décembre de l'année.

Il existe un sens restreint, plus adapté aux études portant sur l'incidence, que nous n'exposerons pas ici, car son développement dépasserait les objectifs de cet article (8).

Quelle que soit la conception d'usage, il existe différents taux.

Taux brut. C'est le taux le plus simple de la fréquence d'un événement au sein d'une population pendant une période donnée. Il exprime une mesure globale, qui ne tient pas compte des facteurs de variation dans la population. On donne les exemples suivants :

- taux de mortalité = nombre de morts / population soumise au risque (toujours incidence, car les morts sont obligatoirement des cas nouveaux) ;
- taux de morbidité = nombre de malades / population soumise au risque (incidence ou prévalence) ;
- taux de létalité = nombre de morts / nombre de malades (incidence).

Taux spécifique. C'est le même, mais dans des sous-populations données ; par exemple par classe d'âge, par sexe, par race, etc.

Un taux spécifique permet d'une part de ne considérer que la population effectivement susceptible ; par exemple le taux d'avortements rapporté aux seules femmes (et non à l'ensemble de la population comportant aussi bien les mâles que les femmes). Il permet d'autre part une mesure plus fine du phénomène, en distinguant des populations qui peuvent avoir un degré de susceptibilité différent ; par exemple le taux de mammites selon le nombre de lactations.

De nombreux autres taux dérivent des précédents, selon l'usage que l'on veut en faire ; par exemple le taux de fécondité, de gestation, etc.

Ratio

C'est le rapport entre les valeurs de deux variables, du type a/b (fig. 1). Dans un premier sens, il vise à comparer un phénomène dans deux groupes différents : il est constitué du rapport des fréquences de deux classes d'une même variable ; par exemple le sex ratio concernant la variable sexe, qui comporte deux classes, mâle et femelle, est le rapport du nombre de mâles au nombre de femmes.

Dans un deuxième sens, élargi, il permet de rapporter un effectif à une grandeur de dimension différente ; par exemple : nombre de vaches au mètre carré, nombre de porcelets nés par truie et par an. Ce type d'indicateur est précieux dans les situations où l'on est dans l'impossibilité d'estimer la population qui devrait figurer au dénominateur (populations transhumantes) : on rapporte alors les cas à des unités de surface.

Nécessité d'une standardisation

Une population est le plus souvent hétérogène. Chaque sous-population peut avoir ses mécanismes propres d'atteinte par une maladie, en tout cas pouvant conduire à des taux spécifiques propres à chacune d'elles. Il est tout à fait possible qu'à mécanismes comparables, deux populations différant seulement par leur composition démographique aient des taux bruts (taux d'ensemble de chacune des populations) différents. Le constat d'une différence significative ne ferait que révéler cette différence démographique. Par conséquent, pour pouvoir comparer les deux populations sur le plan du phénomène épidémiologique lui-même, en vue par exemple de suggérer ou bien de tester des hypothèses explicatives, il faut d'abord neutraliser l'effet des écarts dus aux différences démographiques. Cette opération s'appelle la standardisation, elle permet d'obtenir des effectifs et des taux standardisés, c'est-à-dire comparables par rapport à un critère donné (le plus souvent d'ordre démographique) dont on a gommé les répercussions sur les indicateurs bruts.

Comment récolter l'information ?

La récolte de toute information épidémiologique qui permettra de calculer l'indicateur approprié nécessite la succession de différentes étapes, que l'on peut représenter comme une chaîne, constituée de trois maillons : échantillon / mesure / traitement des données (figure 2). La qualité de la chaîne d'observation épidémiologique est celle du maillon le plus faible.

Sélection des échantillons

Les échantillons sélectionnés doivent être représentatifs de ce qu'ils sont supposés représenter : tout ou partie de la population dont ils sont extraits. Un échantillon est dit représentatif d'une population plus vaste, lorsque à partir

Etapes de la méthode	Qualité (conception, exécution)
1) SELECTIONS des ECHANTILLONS	Représentativité Précision
2) MESURES / INVESTIGATIONS	Fidélité de la mesure
3) TRAITEMENT STATISTIQUE * Présentation des données * Estimation * Comparaison * Extrapolation	* "Règles de statistique descriptive" * Intervalle de confiance * Comparabilité des populations (Standardisation) * Signification statistique * Fidélité de l'indicateur

Figure 2 : Les qualités des trois maillons de la chaîne d'observation épidémiologique (au plan de la conception et de l'exécution).

de lui on peut décrire cette population, non seulement dans sa globalité, mais aussi dans sa diversité. Chaque individu de la population mère a une probabilité non nulle de figurer dans l'échantillon.

Toutefois, l'échantillon n'est jamais la réduction exacte de la population dont il est tiré, comme si une partie de l'information à collecter se trouvait perdue par la procédure d'échantillonnage. En effet, si l'on répète plusieurs fois la même procédure d'échantillonnage, consistant par exemple à tirer 10 animaux d'une ferme comprenant 10 animaux infectés et 90 indemnes, les différents échantillons obtenus seront différents, et avec un nombre variable d'animaux infectés, alors qu'ils proviennent d'un seul et unique ensemble. Cette diversité résulte de ce que l'on appelle la fluctuation, ou l'erreur, d'échantillonnage.

L'erreur d'échantillonnage est d'autant plus forte que la population est hétérogène, et d'autant plus faible que la taille de l'échantillon est importante. Le choix de la méthode de sondage (sondage simple, avec stratification, en grappe, etc.) permet soit pour une taille donnée de l'échantillon de limiter l'erreur d'échantillonnage, soit pour une erreur d'échantillonnage admise ou consentie de réduire les coûts de l'enquête en diminuant la taille de l'échantillon.

Par conséquent, il faut clairement définir la population à laquelle on souhaite extrapoler les résultats du sondage, pour effectuer la sélection de façon adaptée ; par exemple : tous les éleveurs, ou seulement les élevages laitiers, ou les élevages inscrits, ou les plus performants... A défaut, le résultat obtenu pourra être inexact, c'est-à-dire ne pas refléter l'état réel de la population que l'on désirait étudier, mais celui de l'échantillon qu'en pratique les circonstances ont permis d'observer.

Une telle erreur de sélection de l'échantillon par rapport à la population générale à laquelle on désire extrapoler constitue un biais de sélection. Le biais de sélection le plus courant consiste à laisser les individus libres d'entrer ou de ne pas entrer dans l'échantillon : leur motivation (positive ou négative) pour être inclus dans l'échantillon peut être en relation avec le phénomène étudié ; par exemple : sondage sur la fréquence des cheptels infectés d'une maladie faisant appel au volontariat.

Certains pensent que la grande taille d'un échantillon est susceptible de compenser un biais de sélection éventuel, et que, par conséquent, la qualité d'une étude (sous-entendu sa représentativité) serait directement dépendante de la taille de l'échantillon. L'exemple des résultats de consultation électorale publiés progressivement, en fonction de leur disponibilité démontre aisément le contraire : ces résultats portent sur des millions d'électeurs et peuvent fournir une estimation biaisée, et par conséquent différente de l'estimation réalisée sur un

échantillon restreint de 1 500 individus soigneusement sélectionnés... (autrement dit, le nombre ne fait pas la qualité de représentativité d'un échantillon).

Le nombre conditionne la précision d'un résultat : plus l'échantillon est grand, plus l'intervalle de confiance est réduit. Il faut bien distinguer cette qualité de l'exactitude, qui dépend de la représentativité, et donc du mode de sélection de l'échantillon (fig. 3).

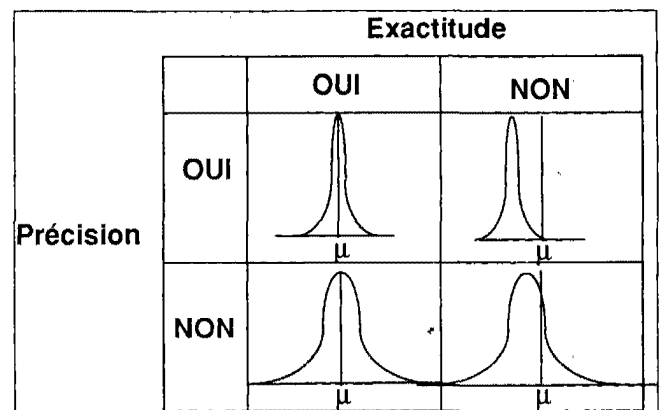


Figure 3 : Relations entre exactitude et précision. L'exactitude résulte de la représentativité de l'échantillon. La précision dépend de la taille de l'échantillon, voire de la stratégie d'échantillonnage (stratification, par exemple).

Mesure ou investigation.

L'information nécessaire aux dénombrements est obtenue par diverses catégories d'observation : prélèvement, questionnaire, mesures variées, etc. Ces différentes observations, quelle qu'en soit la nature, sont des mesures, au sens épidémiologique. Il est capital de se soucier de la qualité de la mesure, qui dépend de plusieurs composantes (1).

Performances

Cette qualité dépend tout d'abord de la qualité de la méthode. Celle-ci dépend, d'une part de ses performances intrinsèques : la sensibilité (probabilité de détection d'un individu réellement affecté) et la spécificité (probabilité de résultat négatif pour un individu réellement non affecté). Elle résulte, d'autre part, de l'effet de la prévalence qui conditionne la probabilité qu'un résultat positif corresponde à un individu réellement atteint de la maladie étudiée (valeur prédictive d'un résultat positif), ou inversement la probabilité qu'un résultat négatif corresponde à un individu réellement indemne de la maladie étudiée (valeur prédictive d'un résultat négatif). Elle affecte la valeur de la prévalence, qui n'est connue que par la fréquence de résultats positifs, dénommée prévalence apparente.

INTRODUCTION

Pour une méthode de qualités connues à l'échelon de l'individu, les performances peuvent être très différentes lorsqu'elle est utilisée en prenant comme unité des cheptels, bien que le test soit appliqué aux animaux : c'est le cas des prophylaxies classiques, comme la tuberculose ou la brucellose pour lesquelles les cheptels font l'objet d'un contrôle de leur état sanitaire à l'aide d'un test réalisé sur tous les animaux des cheptels. Du point de vue de la sensibilité, c'est-à-dire pour détecter un cheptel réellement atteint, il faut et il suffit qu'un seul animal fournisse un résultat positif ; s'agissant de maladies affectant le plus souvent plusieurs animaux dans le troupeau, les éventuelles défaillances de certains animaux sont compensées par le nombre des autres : en définitive, la sensibilité cheptel est toujours très supérieure à la sensibilité individuelle.

On peut prendre l'exemple suivant : pour une sensibilité de 0,90 (très moyenne), la probabilité de ne pas détecter un troupeau comportant un seul animal infecté est de 0,1 ; pour un troupeau comportant deux animaux infectés, elle est de $0,1 \times 0,1 = 0,01$ et pour trois animaux, elle est de 1 pour 1 000, soit une quasi-certitude de détection.

Inversement, en ce qui concerne la spécificité, il faut obligatoirement que tous les animaux du cheptel fournissent un résultat négatif : cette fois le nombre fait en sorte que la probabilité de résultat faussement positif est plus élevée dans un cheptel indemne que pour un animal indemne.

Par exemple, pour une spécificité de 0,99 (excellente), la probabilité de ne pas détecter un animal donnant un résultat faussement positif est de 0,99, pour deux animaux elle est de $0,99^2$, et pour n animaux elle est de $0,99^n$. On peut ainsi vérifier que plus le nombre d'animaux augmente, plus le risque ($= 1 - 0,99^n$) que le cheptel donne un résultat faussement positif augmente : pour 10 animaux, il est de 0,1 pour 40 de 0,33, pour 100 de 0,63, et pour 200 de 0,87.

En résumé, plus le nombre d'animaux testés par cheptel augmente, plus la sensibilité cheptel augmente, la spécificité cheptel diminue, la prévalence apparente augmente et la valeur prédictive négative cheptel augmente (9).

Justesse

La mesure doit rendre compte avec exactitude du phénomène observé, sans le sous-estimer, ni le surestimer d'une façon systématique. Elle doit être distinguée de la précision (figure 3).

La fréquence de résultats positifs observés est la prévalence apparente (P_a). Elle ne doit pas être confondue avec la prévalence réelle (P_r), constituée de la fréquence des individus réellement atteints. La formule suivante donne la relation entre prévalence réelle, prévalence apparente, sensibilité (S_e) et spécificité (S_p) :

$$P_r = \frac{P_a + S_p - 1}{S_e + S_p - 1}$$

Par exemple, pour un test de qualité satisfaisante ($S_e = S_p = 0,95$), et pour une prévalence apparente de 5 p. 100, la prévalence réelle estimée par cette formule est de 1,1 p. 100.

La qualité de la mesure dépend aussi des conditions pratiques (qualité opérationnelle) dans lesquelles cette méthode a été réalisée. En effet, une méthode peut être excellente dans son principe, mais elle peut connaître des variations importantes de la qualité de ses résultats, en fonction de sa fidélité, et de son applicabilité.

Fidélité

Il faut s'assurer que l'instrument de mesure conserve toutes ses qualités lorsque les mesures sont répétées par un même observateur (répétabilité), ou par plusieurs observateurs (reproductibilité), que ce soit au même moment, ou à divers intervalles de temps.

Applicabilité

Le protocole doit pouvoir être mis en oeuvre, exécuté sans que les difficultés soient telles, qu'elles conduisent à un pourcentage de refus trop important, qui par conséquent en constitue une évaluation indirecte. L'enquête peut être trop lourde (questionnaire trop long, informations à collecter trop nombreuses ou difficiles à collecter), nécessiter des prélèvements trop nombreux, ou provoquant des conséquences néfastes sur les animaux (douleur, chute de production).

Son acceptabilité par les enquêteurs doit être optimale : la formation des enquêteurs permet d'expliquer le principe, les objectifs de l'enquête ; elle constitue un préalable à l'acceptabilité. Cette acceptabilité concerne aussi les enquêtés, vis-à-vis de l'intrusion d'un enquêteur dans leur vie privée quotidienne : la présentation de l'enquête, de l'enquêteur, le déroulement de l'enquête elle-même (moment judicieusement choisi, questions claires, non ambiguës, etc.) doivent être soigneusement préparés pour éviter de provoquer des réactions adverses de la part de l'enquêté.

Traitement statistique des données

Le traitement statistique est directement conditionné par la définition des indicateurs, de façon à parvenir aux présentations de résultats souhaités ; par exemple : les présentations de données (tableaux, figures) doivent utiliser les indicateurs et non les données brutes.

Le recours à la statistique permet d'éliminer la subjectivité, d'introduire une rigueur indispensable à la communi-

cation des résultats : tous les résultats doivent être assortis de leur contexte statistique (intervalle de confiance pour les estimations, et jugement "significatif" ou "non significatif" pour les comparaisons).

L'extrapolation des résultats observés sur l'échantillon à la population cible dépend de la fidélité de l'indicateur, en particulier de la qualité de la population qui figure au dénominateur.

La définition d'un indicateur, la conception (et l'exécution) de la chaîne de récolte et traitement d'information nécessaire à son calcul sont bien les caractéristiques fondamentales qui reviennent dans les enquêtes épidémiologiques descriptives.

LES ENQUETES EN EPIDEMIOLOGIE DESCRIPTIVE

Le choix d'une modalité d'enquête descriptive dépend de l'identification des besoins que l'on peut avoir, et de ce qu'une enquête est susceptible d'apporter (les buts qu'elle peut atteindre).

Définition des enquêtes d'épidémiologie descriptive

Les études descriptives ont pour "objet de décrire les caractéristiques d'un phénomène de santé dans une population, son évolution dans le temps, sa répartition, son évolution dans l'espace" (11).

Elles répondent aux questions : Qui ? Quand ? Où ? Comment ? Elles ne répondent pas à la question : Pourquoi ? mais elles contribuent au moins à la poser, voire à suggérer des hypothèses de réponse.

Elles concernent aussi bien les phénomènes pathologiques et les facteurs correspondants que les états et les facteurs de santé (application de mesures thérapeutiques, ou programmes de prévention).

Principe

Elles procèdent par quantification des phénomènes de santé en en déterminant la fréquence selon différentes variables dans les populations, le temps et l'espace. A ce titre, elles font un recours constant à la statistique. Elles ne font que constater des faits. Elles dressent un portrait du phénomène (maladie, ou phénomène de santé quel qu'il soit), ou une série de portraits dans le temps. Ceux-ci peuvent être utilisés pour différentes finalités.

Les besoins visés par les enquêtes d'épidémiologie descriptive

Pour mieux appréhender les enquêtes d'épidémiologie descriptive, il est préférable de commencer par en connaître les finalités, c'est-à-dire qu'une fois le résultat de l'enquête descriptive obtenu, à quoi celui-ci servira-t-il ?

L'information produite par l'épidémiologie descriptive aide à connaître la situation ; elle constitue un outil indispensable dans la prise de décision et dans la planification des mesures sanitaires à entreprendre dans une population donnée. Elle contribue aussi à la formulation d'hypothèses explicatives, qui seront soumises à l'épreuve des faits grâce à des enquêtes analytiques ou à des expérimentations.

Par la connaissance de la situation, l'épidémiologie intervient à différents niveaux de décision dans la conception, la planification, et la gestion de programmes de santé. On donne les exemples suivants :

- définition de priorités d'actions à entreprendre entre différentes maladies constatées sur le terrain ;
- décision du type de prophylaxie, médicale ou sanitaire à appliquer : opportunité de vaccination contre la maladie d'Aujeszký en fonction de la situation sanitaire, dans quels cheptels vacciner, et quels animaux ?
- décisions concernant le pilotage d'un programme de lutte, par exemple contre la tuberculose bovine : 1) repérage des cheptels exposés à un risque particulier, pour y mener une intervention personnalisée ; 2) indication du changement de rythme de tuberculination (actuellement, en France, le rythme est de 1 à 4 ans selon la situation sanitaire) ; 3) identification de la cause de persistance de la tuberculose la plus fréquente dans un département, afin d'en déduire les actions prioritaires ;
- décisions résultant de l'évaluation d'un programme de lutte : les objectifs définis ont-ils été atteints ? Les dépenses consenties ont-elles été convenablement investies ? Le rapport entre les coûts supportés et les résultats est-il conforme aux prévisions ? Est-il acceptable ?

En ce qui concerne l'aide à la formulation d'hypothèses explicatives, l'épidémiologie descriptive intervient à différents niveaux.

Quand une nouvelle maladie survient, il faut commencer par la décrire ; par exemple :

- l'encéphalose hépatique équine qui a sévi dans l'Ouest de la France en 1992 atteint plus particulièrement les juments suitées de 13 ans d'âge en moyenne ;
- de nombreuses erreurs par excès ont été constatées dans le dépistage sérologique de la brucellose bovine en France au cours des campagnes 1990 - 1991 et 1991 - 1992 : elles concernent principalement les élevages à

viande, et ne comportent dans 90 p. 100 des cas qu'un seul animal réagissant.

Il faut ensuite caractériser les phénomènes de manière à pouvoir formuler des hypothèses explicatives ; par exemple :

- 1. l'encéphalose hépatique équine n'est pas de nature infectieuse en raison des caractéristiques de distribution des cas (d'allure non contagieuse) ; elle serait plutôt de nature nutritionnelle ou toxique : dans un lot où une jument fut atteinte, les juments suitées étaient exclusivement à l'herbage, tandis que dans un autre lot de même moyenne d'âge indemne, les juments, également suitées, étaient nourries le soir avec un complément abondant en céréales ;

- 2. la répartition des réactions sérologiques atypiques dans le dépistage de la brucellose bovine est incompatible avec une infection par *Brucella* ; en revanche, les premières tentatives d'isolement menées à partir de bouses ont permis de mettre en évidence *Yersinia*.

But des enquêtes d'épidémiologie descriptive

Le but de l'épidémiologie descriptive est de quantifier, afin d'évaluer la "grandeur" d'un phénomène, sa taille : à dessein, nous n'utilisons pas le mot "importance" qui est très ambigu, car il peut aussi contenir (ou non) tout ce qui concerne les conséquences, en particulier économiques des maladies (dont la mesure fait également partie de l'épidémiologie). Cette quantification, et c'est une des particularités de l'épidémiologie, est rapportée aux populations, au temps et à l'espace.

Distribution des phénomènes de santé dans les populations, le temps et l'espace

Pour répondre à la question "comment procède l'épidémiologie descriptive pour décrire un phénomène de santé dans une population ?" on supposera que toute l'information, toutes les données nécessaires sont présentes : la récolte des informations est envisagée plus loin. La seule question qui demeure par conséquent est de savoir quels sont les résultats auxquels aboutit cette description.

Description dans les populations

A l'aide d'indicateurs appropriés comme le "taux de morbidité" ou le "taux de mortalité", on décrit le phénomène dans une population, en déterminant la valeur de ces indicateurs dans les sous-populations constituées en fonction de caractéristiques de sexe, âge, type d'élevage ou de production, etc.

Répartition et évolution dans le temps

On peut procéder à la représentation de la dynamique d'un phénomène pathologique dans le temps : on peut ainsi en apprécier l'accélération, ou au contraire le ralentissement, ou encore en constater la stagnation. Les variations, si elles existent, peuvent être saisonnières, annuelles, cycliques.

On peut aussi suivre, en fonction du temps, l'importance ou les répercussions d'un phénomène de santé ou des mesures d'intervention correspondantes. Cette information intéresse davantage le gestionnaire, et lui permet, par exemple, d'évaluer les ressources à mobiliser pour l'intervention.

Distribution et évolution dans l'espace

La distribution dans l'espace est étudiée à l'aide de cartes, selon différents modes de représentation (points, codes de densité sous forme d'intensité de couleur par exemple, limites d'extension). Combinée avec le temps, elle permet de suivre la propagation.

Méthodes d'enquêtes descriptives

Les différents types d'enquêtes

On distingue deux types d'enquêtes d'épidémiologie descriptive.

Enquête transversale

Une enquête transversale permet de décrire des phénomènes de santé dans une population sur une période de temps très court ; par exemple : l'étude de la fréquence de la fasciolose dans une région donnée menée sur une période de quinze jours, considérés pour cette enquête comme nécessaires à l'évaluation de la proportion d'animaux atteints. Ce type d'étude convient à la mesure des états, c'est-à-dire de la prévalence.

Enquête longitudinale

Une enquête longitudinale permet de suivre pendant une longue période un phénomène de santé dans un ou plusieurs groupes par des observations périodiques ; par exemple l'étude de la conversion sérologique vis-à-vis de la leucose bovine enzootique dans 50 troupeaux contrôlés tous les mois pendant 3 ans.

Ce type d'étude est nécessaire pour la mesure des changements d'état, c'est-à-dire de l'incidence. Il suppose impérativement de répéter les observations sur les mêmes individus, afin d'être en mesure d'observer le changement d'état. Il réclame des moyens bien plus importants qu'une étude transversale.

Le type de la population qui fait l'objet de l'enquête conditionne aussi le type d'enquête. Si la population est stable, une enquête transversale peut convenir. Si la population est dynamique, soumise à un renouvellement important, le type longitudinal est beaucoup plus adapté.

Les méthodes de sélection des échantillons

L'échantillon doit être représentatif de la population étudiée. Le biais de sélection constitue le défaut dont il faut préserver l'enquête. L'échantillon constitué de façon spontanée, au gré du déterminisme des individus (à tort qualifié de "hasard") est la plus mauvaise des procédures. On va voir les différentes modalités et leurs qualités.

Echantillon empirique

On fixe par avance un nombre de sujets nécessaires par catégorie. Toutes les catégories retenues devront être couvertes par l'échantillonnage, à concurrence du nombre préalablement fixé.

Cette méthode peut être très élaborée, et reposer sur une série de règles précises : il s'agit alors d'un échantillon raisonné. Les multiples sondages réalisés actuellement en sont la meilleure illustration. La démarche consiste à imaginer quelle est la constitution d'un échantillon idéal pour être représentatif, puis à choisir les sujets en conséquence. La question est de savoir comment choisir les individus, ce qui relève du savoir faire des organismes de sondage. A défaut de cette maîtrise, cette modalité d'échantillonnage ne peut être recommandée pour les enquêtes épidémiologiques.

Echantillon aléatoire

Un échantillon aléatoire est un "échantillon défini par un tirage aléatoire (tirage au sort) et dans lequel tous les individus ou toutes les unités ont une certaine probabilité, connue, d'être tirés"(12).

Cette méthode nécessite l'utilisation d'une liste de tous les individus de la population (base de sondage), et le recours à une procédure de sélection aléatoire (table de nombres au hasard, fonction aléatoire d'un ordinateur). Par son principe, elle garantit la qualité de représentativité.

Dans l'échantillon aléatoire simple "chaque individu de la population a la même probabilité d'être tiré" (12). En pratique, il ne donne qu'une image grossière, valable à l'échelon de l'ensemble de la population, dont il donne une valeur moyenne. Si l'on souhaite une image plus fine, par exemple selon différentes catégories d'individus, ou selon la distribution dans l'espace, il faut alors mobiliser des moyens considérables pour satisfaire ce besoin.

Les méthodes de stratification, et de sondage en grappe, permettent d'optimiser le résultat en fonction de res-

sources disponibles, pour une précision souhaitée. On le verra plus loin.

Echantillon pseudo-aléatoire

Le recours au tirage au sort peut être incompatible avec les contraintes pratiques ; par exemple, on ne peut demander un tel tirage au sort sur une chaîne d'abattage. La base de sondage peut aussi ne pas exister. On détermine alors une règle, qui déterminera le choix des sujets inclus dans l'échantillon, et qui doit être indépendante du phénomène étudié.

On peut procéder à un tirage systématique ; par exemple, on décide de prendre un individu sur 10 (ou toute autre nombre) sur un ensemble accessible, et qu'il sera malgré tout possible de classer, même d'une façon sommaire.

Dans une population d'animaux sauvages, dont la structure démographique est connue, on considère que chaque individu a une certaine probabilité d'être capturé, liée à la cohorte à laquelle il appartient.

Il faut éviter une règle qui puisse épouser le rythme cyclique du phénomène étudié. Par exemple, dans un abattoir, choisir un jour par semaine n'est pas une bonne solution, car les arrivages ne sont pas comparables selon les jours de la semaine.

Il faut éviter aussi de constituer des sous-groupes qui pourraient être redondants avec des groupes épidémiologiques ; par exemple en prenant les 5 ou 10 individus survenant à tel moment, on peut de ce fait sélectionner un ensemble provenant d'un même élevage.

Echantillon aléatoire stratifié

Cet échantillon est "réalisé par tirage aléatoire au sein de strates préalablement définies dans la population" (12). Une strate est un sous-ensemble d'une population, plus homogène au regard d'un caractère donné.

On sait, par exemple, que le taux d'élevages infectés dépend de leur taille ; mais, la démographie de ces élevages est telle que si l'on réalise un échantillonnage aléatoire simple, certaines catégories ne seront pas suffisamment représentées dans l'échantillon pour que l'estimation ait une précision satisfaisante. Il peut être nécessaire de sur-représenter certaines, ou d'en sous-représenter d'autres : c'est le principe de la stratification, qui vise à obtenir une meilleure précision d'ensemble, tout en réduisant le nombre de sujets nécessaires. De plus, l'information fournie pour chaque strate est de meilleure qualité.

Echantillon aléatoire en grappes

Les échantillons aléatoires précédents rencontrent des obstacles pratiques qui peuvent être insurmontables : il faut pouvoir accéder à chaque individu sélectionné à partir de la base de sondage, ce qui peut demander des

INTRODUCTION

efforts considérables, en particulier des déplacements coûteux en temps et en argent. Ces inconvénients sont résolus lorsque les individus sont regroupés en unités collectives, comme des logements, des villages, ou encore une clientèle vétérinaire.

L'échantillonnage consiste à tirer au sort un échantillon de ces unités collectives, et à étudier tous les individus qui y sont contenus, d'où le terme de "grappe".

Pour que ce type de sondage soit efficace, il faut que les grappes soient relativement homogènes entre elles, mais qu'elles couvrent une large dispersion des valeurs pour chacune d'elles : ainsi, malgré un faible nombre de grappes tirées, l'image sera représentative de la population.

Inversement, si les grappes sont très différentes entre elles, par exemple parce qu'il existe des zones atteintes, et d'autres non atteintes, alors que la variabilité à l'intérieur de ces grappes est faible, le processus de tirage des grappes peut fournir une mauvaise image de la population.

Echantillon aléatoire à plusieurs degrés

Un échantillon à plusieurs degrés résulte d'un échantillonnage réalisé en plusieurs étapes.

Les unités de sondage sont organisées en différents niveaux hiérarchiques : par exemple, les individus sont regroupés en élevages, eux-mêmes regroupés en villages. La procédure de sélection consiste à tirer au sort un nombre donné d'unités correspondant à chacun des niveaux : un certain nombre de villages, dans lesquels on sélectionne un certain nombre d'élevages, dans lesquels on sélectionne un certain nombre d'individus.

Les unités de la première étape sont encore appelées "unités primaires", celles de la deuxième étape "unités secondaires", et ainsi de suite. Le tirage des unités primaires est le premier degré de l'échantillonnage, celui des unités secondaires le deuxième degré, etc.

Cette méthode réunit les avantages du sondage en grappes, qui permet la concentration des moyens. On n'a besoin que de la liste des unités de sondages soumise au tirage, et non de tous les individus.

L'avantage de ce type de sondage provient de l'économie réalisée dans les grappes, qui en permettant d'étudier un plus grand nombre de grappes, conduit à espérer une meilleure précision, afin de compenser une diversité éventuelle entre les grappes qui pourrait être trop grande.

Nombre de sujets nécessaire

On n'évoquera que brièvement le principe de la détermination du nombre de sujets nécessaire sur l'exemple de l'échantillon aléatoire simple, en renvoyant à des articles spécialisés pour les autres modalités d'échantillonnage,

dont certaines peuvent nécessiter l'aide d'un statisticien confirmé (4, 5, 6).

On veut estimer un pourcentage, avec un intervalle de confiance donné. Le nombre de sujets nécessaire dépend de la taille de la population dont l'échantillon provient, de la valeur de la prévalence à estimer, et la précision souhaitée.

Plus le pourcentage à estimer est petit, plus le nombre de sujets nécessaire est élevé. Il faut par conséquent en faire l'hypothèse préalable, sur la base d'informations disponibles avant l'enquête. Si le pourcentage trouvé est plus élevé, la précision obtenue est supérieure à celle qui était souhaitée, et on a dépensé des moyens de façon inutile. Si le pourcentage est inférieur à la valeur théorique, la précision est moindre : il reste à vérifier si elle est encore compatible avec les décisions qui doivent en résulter.

L'intervalle de confiance, ou la précision, est à déterminer en fonction de considérations pratiques, de l'usage qui sera fait de l'information ainsi produite. Par exemple, on peut être amené à choisir entre deux stratégies de lutte (sanitaire ou médicale) en fonction de seuils d'infection, qu'il importe de pouvoir nettement distinguer, grâce à une précision suffisante. Ou bien, on peut avoir besoin d'estimer l'importance de moyens à mettre en oeuvre : ce sont alors les contraintes liées à ces moyens qui détermineront la précision nécessaire.

Cette précision peut être rapportée à la valeur de la prévalence cherchée, sous forme d'une précision relative, qui facilite considérablement la détermination du nombre nécessaire, par la simple consultation de tableaux, ou d'abaques (4). Ainsi, une précision absolue de 5 p. 100, n'a pas les mêmes répercussions sur la précision d'un pourcentage à estimer qui serait soit de 50 p. 100, soit de 10 p. 100. Dans le premier cas, la précision relative est de 10 p. 100, tandis que dans le deuxième elle est de 50 p. 100.

Les modalités d'évaluation de ces enquêtes seront étudiées dans la partie "Discussion".

LES ENQUETES EN EPIDEMIOLOGIE ANALYTIQUE

Définition

Les enquêtes analytiques ont pour objectif de "déterminer le rôle d'un ou de plusieurs facteurs dans l'étiologie d'une maladie" (12). Cette partie de l'épidémiologie vise à répondre à la question : "pourquoi ?", à l'aide de démonstrations appropriées. A minima, elle contribue aussi à la formulation d'hypothèses, plus fortes que celles de l'épidémiologie descriptive.

Principe

Le but des enquêtes analytiques est de soumettre des hypothèses préalablement formulées à l'épreuve des faits. Leurs méthodes se basent sur des comparaisons entre des groupes. Ces comparaisons sont accomplies en estimant et en testant le degré d'association entre un facteur supposé causal et son effet (la maladie).

La simple constatation d'une association statistique ne suffit pas : il faut aussi mesurer l'intensité de cette relation, et surtout soumettre la nature de cette relation à l'épreuve des faits : est-elle fortuite, associée en fait à un autre facteur, ou véritablement causale ? Tout l'esprit de cette démarche est d'ordre essentiellement méthodologique.

Commentaires

Le mot "analytique" est ambigu. On peut en effet procéder à une analyse statistique de données descriptives, sans envisager de but explicatif. Le terme même d'analyse renvoie à un processus de décomposition en éléments qui ne correspond pas à la démarche explicative par modélisation qui est plutôt synthétique. Il serait préférable de lever toute ambiguïté en utilisant le terme d'épidémiologie explicative ou étiologique pour ce qui concerne l'étude des causes.

Les études épidémiologiques à vocation explicative sont riches de deux démarches scientifiques complémentaires, par expérimentation (essais épidémiologiques), et par observation (enquêtes exposés / non-exposés, cas / témoins, et de prévalence), auxquelles nous limiterons notre exposé.

On ne peut tirer de conclusion explicative qu'au travers d'une interprétation qui prend en compte non seulement la qualité du protocole d'observation, mais aussi les conditions propres aux principes de la logique d'inférence par déduction, appliqués à l'épidémiologie. C'est pourquoi, pour plus de facilité dans la présentation, on exposera tout d'abord le principe de l'étude de la relation de cause à effet au travers du modèle fondamental de l'expérimentation. Ensuite, les trois méthodes essentielles, enquêtes cas/ témoins, enquêtes exposés / non-exposés et enquêtes de prévalence seront présentées. A partir de ces trois méthodes de base dérivent une grande diversité de types d'enquêtes. Pour finir, les conditions d'interprétation de ces diverses méthodes, permettant de conclure sur un plan explicatif, seront étudiées.

Les besoins visés par les enquêtes analytiques

Les enquêtes visent tout d'abord à établir des connaissances permettant de comprendre les mécanismes épidémiologiques des maladies. Leur interprétation n'est

pas toujours facile, et de nombreux exemples de médecine humaine (relation tabac-cancer bronchique) illustrent cette difficulté. En effet, de nombreuses différences autres que le facteur étudié, de surcroît liées au facteur lui-même, permettent de distinguer les populations atteintes et non atteintes ; on ne peut donc avoir la certitude de la responsabilité causale de ce facteur incriminé, à la différence des études expérimentales.

Ces difficultés peuvent conduire à viser des objectifs moins ambitieux, mais plus pragmatiques. Les enquêtes peuvent chercher à mettre en évidence des facteurs permettant de prédire, plutôt que d'expliquer. Les informations correspondantes sont utiles pour sélectionner les populations concernées, afin de prendre les mesures de prévention adaptées.

Enfin, toutes les méthodes utilisées dans la lutte contre les maladies peuvent faire l'objet d'étude : ce sont les études d'intervention, qui à la différence des études d'évaluation ont une structure de type analytique. On peut en distinguer différents niveaux, selon le degré de qualité du contexte de réalisation. Par exemple, l'étude d'une méthode de dépistage sera étudiée tout d'abord dans un laboratoire parfaitement équipé, possédant un standard de qualité d'exécution maximal, en comparaison à une méthode de référence. L'étude peut ensuite être menée dans le contexte des laboratoires qui auront à appliquer la méthode, dans les conditions de la pratique. Pour pouvoir rapporter les différences constatées à la seule méthode et à ses conditions locales d'application, et non à d'autres éléments indépendants appartenant au contexte, la structure de l'étude doit être de type analytique.

Principe de l'étude de la relation de cause à effet. Les études expérimentales

Le problème de la cause et de l'effet

Ce n'est qu'*a posteriori* que l'on peut attribuer à tel événement (A) le rôle de cause et à tel autre (B) celui d'effet. Pour y parvenir, on doit vérifier, par l'expérience des faits constatés, un ensemble de vérités logiques :

- A précède B,
- l'association entre A et B est "suffisamment" forte,
- aucun autre événement que A n'est susceptible d'intervenir pour perturber l'étude.

Cette relation, telle qu'elle est présentée, est simpliste. Elle n'envisage le rôle que d'un seul facteur. Elle néglige toutes les possibilités d'association et d'interactions de plusieurs facteurs. Elle permet toutefois d'expliquer plus simplement la problématique de l'étude de la relation de cause à effet. A partir de ce modèle, il est possible de transposer à des modèles plus complexes.

Le modèle expérimental de Claude Bernard

Les conditions évoquées précédemment sont réunies dans ce prodige logique que constitue le modèle expérimental formalisé par Claude Bernard.

Principe

Deux lots au moins sont comparés : l'un est soumis (par l'expérimentateur) à une manoeuvre expérimentale (A+), l'autre subit exactement les mêmes manoeuvres, au facteur étudié près (A-), et sert de témoin (seul le facteur étudié est susceptible d'intervenir). On compare l'apparition de l'effet (B) dans chacun des lots.

Les conditions tenant au protocole (figure 4)

Conditions tenant au protocole	Conditions tenant à la logique du raisonnement scientifique
1/ Formulation préalable d'une hypothèse	1/ Respect du cycle induction - déduction
2/ Maîtrise du facteur étudié	
3/ Comparabilité : initiale finale	2/ Stratégie de réfutation

Figure 4 : Les conditions permettant la conclusion de cause à effet à partir d'une expérimentation.

La formulation préalable d'une hypothèse qui prédit les résultats escomptés est indispensable ; elle permet d'élaborer le protocole, dans ses détails d'observation, mais aussi du point de vue de l'analyse statistique qui sera utilisée pour apprécier la force de la relation.

Les conditions expérimentales doivent permettre de garantir que le facteur A précède effectivement l'apparition de l'effet B, et de maîtriser les facteurs externes qui seraient susceptibles d'interférer sur l'un et/ou l'autre groupes observés.

Le tirage au sort des individus qui seront affectés à chacun des lots, à partir d'une population homogène, apporte une garantie déterminante de leur comparabilité initiale. Les conditions d'entretien et d'observation des sujets d'expérience doivent être totalement comparables pour les lots, de façon à maintenir leur comparabilité jusqu'à la fin de l'expérience.

Conditions tenant à la logique du raisonnement scientifique

Respect du cycle logique d'induction / déduction. La logique de la connaissance scientifique procède par cycles d'induction des hypothèses par des observations initiales, et de déduction par la mise à l'épreuve des faits de ces mêmes hypothèses par de nouvelles observations. On ne peut, sur un même jeu de données, procéder à ces deux phases. La formulation de l'hypothèse

explicative est donc préalable à l'élaboration du protocole explicatif.

Respect de la stratégie de réfutation d'hypothèse. La stratégie d'interprétation scientifique est *a priori* celle de la réfutation, c'est-à-dire du rejet de l'hypothèse soumise à l'épreuve. Ce n'est que par un ensemble de déductions que l'on est conduit à admettre les faits nouveaux. On ne peut donc pas confirmer directement une hypothèse.

Par exemple, l'isolement d'une Salmonelle de même sérotype dans deux élevages possédant des liens épidémiologiques ne permet pas d'affirmer que l'un est à l'origine de la contamination de l'autre. Ils ont pu être tous deux contaminés par une même source extérieure. En revanche, la mise en évidence d'une différence minime, par exemple d'ordre génétique, permettant d'affirmer la non identité des souches est suffisante pour rejeter l'hypothèse d'un lien épidémiologique de contamination entre les deux élevages. Ainsi, les faits permettent d'infirmer une hypothèse, jamais de la confirmer directement. Ce n'est que l'interprétation des résultats qui conduit à admettre sa confirmation comme plausible, dans la limite des possibilités de réfutation.

Cette stratégie de la connaissance scientifique pose de délicats problèmes d'interprétation pour les études épidémiologiques, en raison de leur particularité qui est de ne pas satisfaire l'ensemble des critères précédents.

Les études d'observation

L'épidémiologie dispose, en complément des études par expérimentation dont on vient de voir le principe, d'études qui ne réunissent pas l'ensemble des composantes du modèle expérimental, en particulier le facteur n'est jamais maîtrisé par l'observateur. C'est pourquoi elles sont qualifiées d'études d'observation, car les événements se déroulent indépendamment de la volonté du chercheur, qui ne peut ni les provoquer, ni en maîtriser le cours, mais seulement les observer. Toutefois, il peut contrôler les conditions de l'observation.

Pour étudier l'association entre un facteur supposé causal (F) et la maladie (M), différents types d'enquêtes sont possibles, se ramenant à 3 modes d'échantillonnage :

- le chercheur détermine les lots selon le facteur, et il observe l'apparition de la maladie, il s'agit des enquêtes exposés / non-exposés ;
- le chercheur détermine les lots selon l'état par rapport à la maladie, et il mesure la fréquence du facteur, ce sont les enquêtes cas / témoins ;
- enfin, le chercheur observe une seule population, dans laquelle les individus se distribuent selon leur état vis-à-vis du facteur et de la maladie, ce sont les enquêtes de prévalence.

Enquête exposés/non-exposés (fig. 5)

Au début de l'enquête, l'information concernant le facteur d'exposition étudié est connue sur tous les individus, avant que la maladie n'apparaisse. Deux lots sont ainsi constitués, l'un est spontanément "exposé" au facteur, l'autre non. La population à risque (exposée et non-exposée) est suivie pendant une période de temps donnée. L'enquête exposés / non-exposés permet ainsi de recenser les cas incidents apparus pendant cette période de suivi, chez les exposés et chez les non-exposés.

Ce dispositif ressemble beaucoup à l'essai épidémiologique : il se déroule de manière "prospective" dans le temps, d'où la dénomination aussi d'enquêtes "prospectives" que l'on donne aussi aux enquêtes exposés / non-exposés. Mais il s'en distingue par le fait que l'observateur ne maîtrise pas les conditions d'exposition au facteur causal étudié : les sujets sont spontanément exposés ou non, et l'observateur n'intervient pas dans ce choix ; ce qui veut dire que l'on n'a pas la certitude équivalente à celle des conditions expérimentales en ce qui concerne les modalités d'exposition ou non au facteur, l'intensité, la durée, etc.

La comparabilité peut être recherchée *a priori*, c'est-à-dire avant le début de l'observation, à l'aide d'une procédure d'appariement : on choisit les témoins non-exposés de façon qu'ils ressemblent au mieux aux sujets exposés, par exemple du point de vue du sexe, de l'âge, du type d'élevage, etc. Elle peut aussi être établie *a posteriori*, à l'aide de procédures statistiques appropriées (ajustement), pour autant que les effectifs soient en nombre suffisant. Les difficultés sont d'obtenir un échantillon de sujets exposés qui soient véritablement représentatifs des exposés de la population générale dont ils proviennent, et que les lots restent comparables entre eux.

Cependant, dans le cas d'une maladie rare ou d'une période de latence (délai entre l'exposition et l'apparition de la maladie) trop longue, l'étude prospective nécessite un nombre de sujets trop élevé et des périodes d'observation très longues. Ces deux paramètres sont difficiles à gérer sur le terrain (abandon, vente d'animaux, cessation d'activité, etc.).

La relation est mesurée par le rapport des taux d'incidence (risque) chez les exposés et chez les non-exposés : c'est le risque relatif (RR) (voir fig. 7).

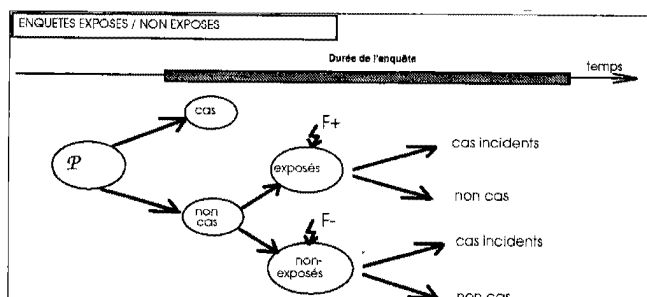


Figure 5 : Enquêtes exposés/non-exposés. P = population de base ; F = facteur étudié.

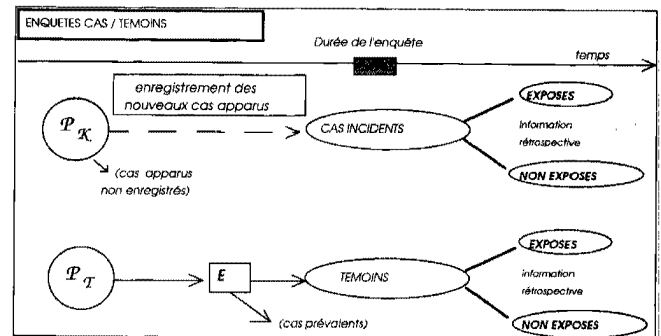


Figure 6 : Enquêtes cas/témoins. P_K = population à partir de laquelle les nouveaux cas enregistrés sont apparus ; P_T = population à partir de laquelle on tire les témoins ; E = échantillonnage (les cas prévalents sont exclus).

Enquête cas / témoins (fig. 6)

Un lot est constitué de sujets atteints ("cas"), un autre de sujets non malades dénommés "témoins" ou "non-cas". La comparaison est effectuée à l'aide d'informations sur l'exposition antérieure à un ou plusieurs facteurs. Cette information est collectée au moment de l'enquête, alors que les événements ont déjà eu lieu. Cette démarche, qui consiste à remonter dans le temps pour reconstituer les informations, fait aussi qualifier cette catégorie de "rétrospective".

Il faut disposer d'un système de suivi de l'état sanitaire de la population qui permette d'enregistrer les cas incidents, au fur et à mesure de leur apparition. Les témoins sont choisis à partir de la population générale, une fois que tous les cas incidents nécessaires à l'enquête ont été détectés, et après élimination des cas prévalents.

Les deux populations (P_K et P_T) ne sont pas systématiquement équivalentes. En effet, le système de recensement des cas peut exclure une catégorie de la population, et le délai entre la détermination des cas et des témoins peut rendre les deux populations non comparables.

Pour pouvoir faire de l'inférence causale, il faut vérifier ou supposer que les non-cas (témoins) tirés de la population (P_T) sont représentatifs de la population (P_K) à partir de laquelle les cas se sont développés, c'est-à-dire que les deux populations sont comparables. Afin d'assurer la comparabilité, des techniques d'appariement peuvent être utilisées.

Dans le cas de la listériose humaine, les patients sont appariés avec des témoins, sur des critères par exemple de sexe, âge, catégorie socio-professionnelle, race. Par exemple, pour un cas survenu chez une femme de 32 ans, institutrice, de race blanche, on apparie (au moins) un témoin de même sexe, âgé de 25 à 35 ans, travaillant dans le secteur tertiaire, de même race.

La mesure de l'intensité du lien entre le facteur étudié et la maladie ne peut être effectuée de la même manière que dans les études exposés/ non-exposés, en raison de la différence fondamentale de constitution des populations. Les enquêtes exposés/non-exposés (figure 7) permettent de mesurer l'incidence, en raison de leur caractère prospectif. Les enquêtes cas/témoins (figure 8) sont constituées de manière totalement différente, représentée dans la figure par une orientation perpendiculaire du tableau d'analyse par rapport à celui des enquêtes exposés/non-exposés : on mesure la fréquence d'exposition au facteur étudié. Cette démarche, par nature rétrospective, interdit tout calcul d'incidence : la somme d'effectifs en lignes, nécessaire à ce calcul, n'aurait en effet aucun sens. L'odds ratio, tel qu'il est défini dans le tableau, constitue la mesure recherchée du lien entre les variables. Un odds est le rapport d'une probabilité à son complément ($p/(1 - p)$). L'odds ratio est un rapport de deux odds (12), odds chez les sujets exposés au facteur rapporté à l'odds chez les non-exposés. L'odds ratio mesure l'intensité ou la force liant un facteur de risque à un état pathologique. La liaison est d'autant plus forte que la valeur de l'odds ratio est plus élevée. Dans le cas de maladies rares, il constitue aussi une approximation du risque relatif.

Dans certaines études, il n'existe pas de système de déclaration de la maladie permettant de répertorier les cas incidents. Le recrutement des cas et des non-cas se fait alors à partir de la population générale, et les cas deviendront des cas prévalents.

Enquête de prévalence

L'échantillon est tiré au sort à partir d'une population cible définie par l'investigateur. Comme le montre la figure 9, dans ce type d'enquête, les cas (malades) appartiennent à la population des cas prévalents. Ce type d'enquête ne prenant pas en compte le temps est aussi qualifié d'enquête transversale.

Après la sélection des animaux ou groupes d'animaux (bandes, élevages, etc.), tous les individus sont examinés ou observés, leur statut d'exposés ou non-exposés est déterminé et d'autres variables sont parfois relevées.

Dans certains cas, l'échantillon n'est pas aléatoire. Il faut noter que sans échantillonnage aléatoire, une étude de prévalence a une valeur très limitée si elle a comme objectif de déduire la fréquence de la maladie ou autres caractéristiques d'une population et présente ainsi de très sérieuses limitations pour faire de l'inférence causale. Elles peuvent malgré tout servir dans une phase de prospection initiale à la proposition d'hypothèses.

Quand la maladie est rare, ce type d'étude reste peu puissant pour détecter d'éventuelles relations entre un facteur et la maladie.

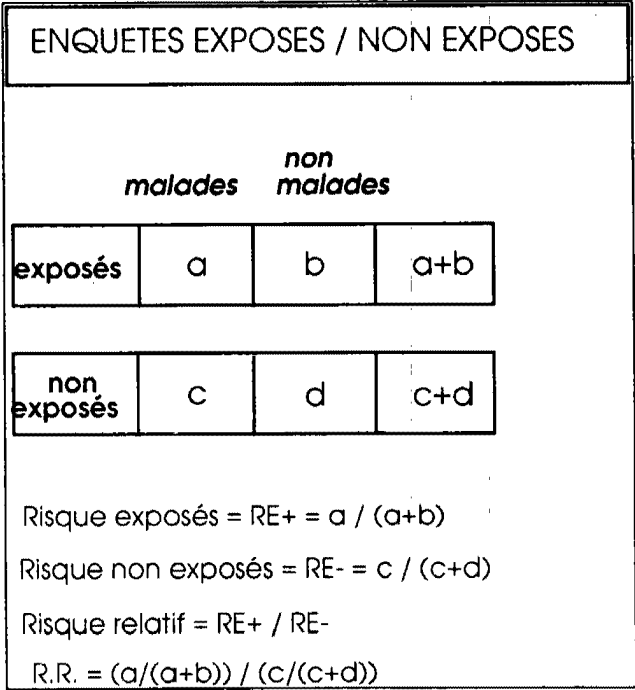


Figure 7 : Les mesures de l'intensité des relations dans les enquêtes exposés/non-exposés.

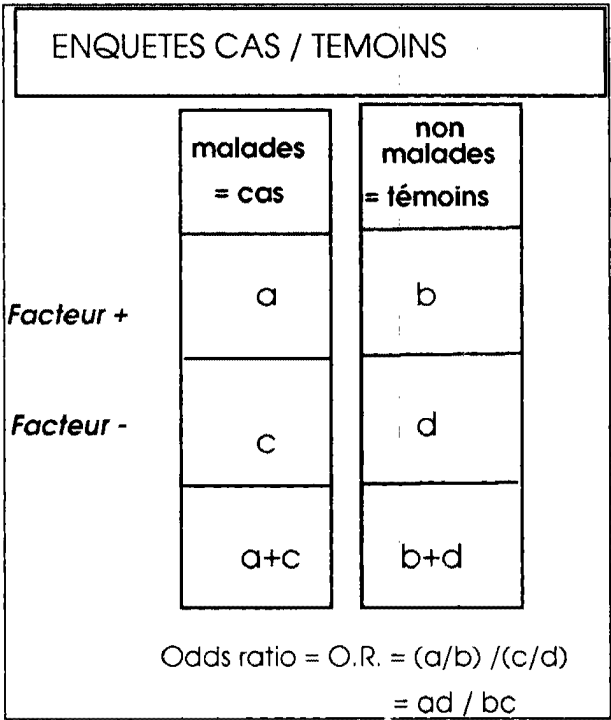


Figure 8 : Les mesures de l'intensité des relations dans les enquêtes cas/témoins.

Détermination du nombre de sujets nécessaire

Comme dans la partie consacrée aux enquêtes descriptives, nous ne présenterons que le principe de la détermination du nombre de sujets nécessaires.

Le problème se résume à la comparaison de deux pourcentages (P_1 et P_2) de la maladie chez les exposés et les non-exposés (dans le cas d'une enquête prospective), ou pourcentages d'exposés chez les cas et les témoins (enquête rétrospective). Le test statistique consiste à comparer la différence observée entre les échantillons à la différence maximale qui peut être due au hasard résultant de fluctuations d'échantillonnage (fig. 10). Si la différence observée est supérieure à cette différence théorique, elle est considérée comme significative. Cette interprétation est donnée avec un risque résiduel d'erreur (α) que le hasard puisse malgré tout expliquer cette différence. Inversement, si la différence observée est inférieure à la différence théorique, elle est considérée comme non significative ; cette dernière interprétation est donnée avec un autre risque d'erreur (β) que l'écart en fait ne soit pas dû au hasard, mais bien à une différence substantielle entre les deux populations. Par conséquent, l'aptitude du test à révéler cette différence réelle est sa puissance, dont la valeur est le complément à l'unité de l'erreur β .

Le nombre de sujets nécessaire pour révéler la différence entre ces deux pourcentages dépend donc (11) :

- de l'écart réel (qui à la conception du protocole ne peut qu'être posé en hypothèse) ; plus il est élevé, et plus le nombre de sujets nécessaire est faible ;
- du risque d'erreur α consenti : plus il est faible, plus le nombre de sujets devra être élevé ;
- de la puissance du test ($1 - \beta$) : plus elle est élevée, plus le nombre de sujets devra être élevé.

Interprétation d'une relation de cause à effet

L'interprétation des enquêtes épidémiologiques explicatives doit toujours être prudente, en raison de l'absence de maîtrise des conditions d'observation (en particulier du degré d'exposition aux différents facteurs causaux étudiés pour les études d'observation) à la différence des expérimentations. Les conclusions ne peuvent pas être portées directement en termes de relations causales entre les facteurs de risque observés et la maladie. En effet, une association peut être fallacieuse, causale ou non causale.

L'interprétation causale

Une fois la validité interne établie (voir "Discussion"), on peut envisager d'interpréter les résultats dans le sens

recherché de la relation de cause à effet. Pour pouvoir conclure sur une telle relation, il est admis dans la communauté scientifique qu'il faut et il suffit que les conditions méthodologiques du modèle expérimental soit strictement respectées : les contraintes méthodologiques imposées correspondent en effet à l'ensemble des conditions logiques permettant de conclure sur le plan causal. La discussion porte habituellement sur l'adéquation entre le modèles expérimental et le problème que l'on veut résoudre, et les possibilités de généralisation à partir des observations effectuées.

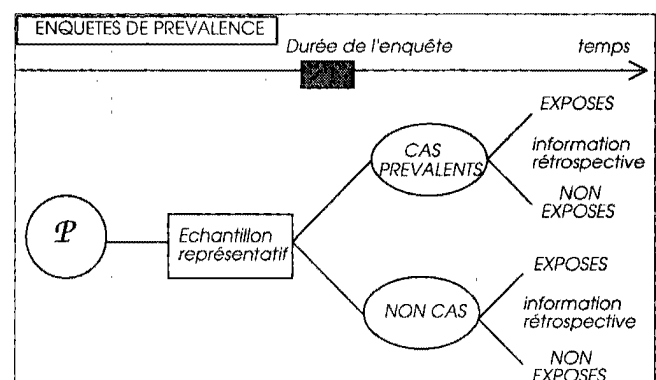


Figure 9 : Enquête de prévalence.

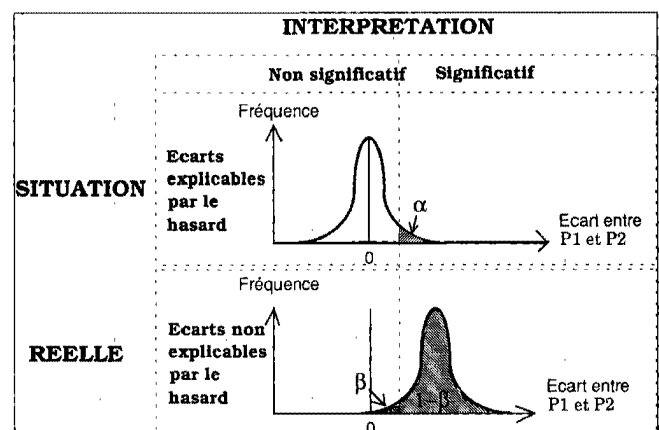


Figure 10 : Comparaison de deux pourcentages : les risques d'erreur dans l'interprétation des résultats du test statistique, en fonction de la situation réelle.

Quand les écarts sont explicables par le hasard des fluctuations d'échantillonnage, les écarts entre les pourcentages P_1 et P_2 sont distribués normalement autour d'une moyenne nulle. Si un écart observé est suffisamment différent de 0, le résultat du test est considéré "significatif". Dans le cas où les écarts sont dus au hasard, cette interprétation est donnée avec un risque d'erreur α . Dans le cas où les écarts ne sont pas dus au hasard, l'interprétation "résultat non significatif" est donnée avec un risque d'erreur inverse (β) ; l'aptitude d'un test à donner la bonne interprétation est sa puissance ($1 - \beta$).

INTRODUCTION

En épidémiologie, par principe, les conditions intrinsèques du modèle expérimental sont mal, voire pas du tout, respectées, au moins pour certaines d'entre elles. L'épidémiologiste doit par conséquent apprendre à gérer par lui-même les conditions logiques permettant une conclusion de cause à effet : les différentes propositions de telles règles, que nous allons évoquer, révèlent la simplicité et l'extrême efficacité du modèle expérimental.

On retiendra ici les "critères" proposés par HILL (1965, *in* (7)).

La séquence dans le temps : la cause précède l'effet

Cela peut paraître une évidence. Mais il n'est pas toujours facile de pouvoir garantir que le facteur étudié précède effectivement l'apparition de la maladie. Par exemple, il faut tenir compte du délai de latence, il faut pouvoir dater les événements. Une simple réaction sérologique positive peut ne pas être suffisante, en raison de la durée de persistance des anticorps (toute la vie de l'individu pour certaines).

La force de l'association statistique : cause et effet sont fortement corrélés

La force ne veut pas dire degré de significativité, mais renvoie à l'intensité du lien entre les variables, risque relatif et odds ratio selon le cas, qui eux doivent avoir une valeur significative.

L'essai randomisé : les conditions d'observation doivent exclure les biais

Ce critère constitue la reconnaissance la plus claire des limites des études d'observation. Mais il est vrai que la randomisation constitue un moyen de garantir la comparabilité au moins initiale des lots à comparer.

La reproductibilité des résultats sur d'autres échantillons : on doit aboutir aux mêmes conclusions

Ce critère est très général, mais il a d'autant plus d'importance en épidémiologie que les sujets observés ne sont pas homogènes : les résultats peuvent être particuliers à la population observée, et ne pas être extrapolables à d'autres populations.

Le gradient biologique

L'existence d'une relation dose-effet constitue un bon indicateur de la solidité des conclusions.

La plausibilité et la cohérence des observations

Les conclusions doivent être en accord avec les connaissances antérieures, ou les modèles (à moins que l'on ne cherche à en démontrer un nouveau), ou les théories.

Les dogmes, en revanche, sont faits pour être soumis à la critique scientifique aussi souvent que possible, pour autant que l'on soit capable de les identifier parmi l'ensemble de connaissances scientifiques communément admises. Ce critère est très conservateur, mais il est robuste : une connaissance nouvelle a moins de risque d'être artefactuelle si elle ne remet pas en cause les fondements de la connaissance (l'exemple de la mémoire de l'eau de BENVENISTE est encore dans nos mémoires (3)).

Les éléments précédents du raisonnement causal ne font que permettre d'apprécier la validité de la réponse que constituent les résultats de l'étude à la question posée au travers du protocole, le problème à résoudre étant celui de la validité de l'hypothèse à tester.

DISCUSSION

La présentation des méthodes serait incomplète sans l'évocation de la stratégie de choix entre les différentes modalités d'enquête, et de l'évaluation de la validité des résultats.

Choix stratégiques dans les enquêtes épidémiologiques

Descriptif, explicatif ?

L'exposé précédent a opposé les enquêtes descriptives et analytiques. Cette distinction est effectivement importante du point de vue de la satisfaction des critères de qualité correspondant : représentativité pour les enquêtes descriptives, et comparabilité, contrôle des tiers facteurs pour les enquêtes analytiques. Les buts visés sont très différents.

Mais, en pratique, les besoins sont souvent complexes, et peuvent emprunter à la fois à l'un et à l'autre domaine. Comment les concilier, tout en préservant ce qu'il est indispensable de rigueur pour ne pas aboutir à des résultats sans aucune valeur ?

Il faut reconnaître que la présentation faite a volontairement simplifié, pour les besoins de la clarté, ce qui en pratique est beaucoup plus nuancé.

Ainsi, sur la base d'un jeu de données purement descriptives, il n'est pas interdit d'utiliser des méthodes de type analytique, et ainsi de dépasser les buts descriptifs, en vue de proposer des hypothèses explicatives. Celles-ci auront une bien meilleure validité que si l'on s'était borné à la simple constatation d'associations. Mais, ces constatations ne seront que des hypothèses, et c'est bien là la frontière avec le domaine explicatif véritable, qui lui commence après la formulation de l'hypothèse.

D'un point de vue réaliste, encore une fois, force est aussi d'admettre que les contraintes de terrain peuvent être telles que même un protocole proprement analytique peut voir sa valeur considérablement altérée, au point que la force des conclusions peut être du même ordre que celles provenant d'une étude descriptive, dont on aurait exploité les données dans un sens analytique. Que vaut un tel protocole, si les tiers facteurs ne peuvent être contrôlés, si l'antériorité du facteur étudié sur l'effet n'est pas vérifiée, ou si les conditions de comparabilité des lots ne sont pas satisfaites ?

En définitive, la démarche scientifique d'explication suppose dans son essence principale une remise en jeu des hypothèses par un apport de nouveaux faits. Ceux-ci peuvent provenir de différents types de protocole, qu'il soit expérimental, analytique, ou autre. L'exemple de salmonellose commune à deux élevages, a montré que des constatations descriptives peuvent permettre de rejeter une hypothèse explicative.

Autrement dit, mieux vaut réaliser de bonnes études descriptives dont les résultats seront exploitables, et mieux vaut en associer, en combiner plusieurs selon une stratégie de raisonnement bien construite, que de mener de mauvaises études analytiques qui ne permettraient aucune conclusion.

On vient, en quelque sorte, de caractériser les particularités méthodologiques d'un grand nombre d'études épidémiologiques qui ont été dénommées "écopathologiques". Celles-ci visent à étudier la pathologie dans ses rapports avec le milieu, par exemple, l'infécondité de la vache laitière, en fonction de l'alimentation, de la conduite d'élevage, etc., dans une approche systémique.

Bien souvent, les protocoles écopathologiques ne peuvent être clairement rattachés à l'un des types développés dans cet article, au point qu'il a semblé nécessaire à certains de créer un domaine à part, pour justifier cette originalité. En effet, des biais initiaux importants obèrent souvent toute velléité de représentativité, car le recrutement des propriétaires dépend de leur bonne motivation. De même, les lots à comparer sont constitués spontanément pour les mêmes raisons, et la comparabilité doit être assurée *a posteriori*, ce qui nécessite de grands effectifs. Les méthodes statistiques utilisées pendant longtemps étaient seulement descriptives.

Malgré ces limites, ces enquêtes ont contribué à la connaissance de différentes maladies. En effet, elles consistent en la succession de campagnes d'enquêtes, en partant d'un état de prospection en quête d'hypothèse, pour éprouver ces hypothèses par de nouveaux jeux de données sur des échantillons plus importants. Elles peuvent aussi les soumettre à l'épreuve des faits, par la mise en oeuvre de mesures de lutte adaptées à chaque cas d'élevages, sur la base des prédictions résultant du traitement statistique des observations précédentes. Même si des modalités pratiques différentes ont pu être décrites selon les équipes, dans leur ensemble, elles res-

pectent le principe fondamental du cycle induction-déduction, qui est une des bases essentielles du raisonnement scientifique.

Il faut toutefois souligner qu'elles demeurent dans un domaine résolument pragmatique, et qu'à défaut de pouvoir respecter les critères de qualité nécessaires, elles ne peuvent pas être considérées comme des études de méthodologie explicative au sens strict. On n'ira pas plus loin dans ce débat d'une brûlante actualité internationale, puisque régulièrement des congrès abordent cette question épistémologique du rapport entre l'explication et la décision (10).

Au total, le choix entre descriptif et explicatif dépend de la balance entre les besoins (quelle information est nécessaire pour quelles décisions), et les ressources, au sens le plus large. De plus, il faut aussi tenir compte du fait que la description n'est jamais suffisante ou trop riche, et que sur sa base, il est possible de prendre quantité de décisions, pourvu que l'interprétation soit conduite en connaissance des limites des résultats.

Enfin, on peut considérer qu'en dehors de protocoles qui peuvent être très clairement classés comme descriptifs ou analytiques, il est possible d'envisager une espèce de continuum passant de l'un à l'autre de ces aspects au travers d'une grande variété de protocoles (8).

Enquête longitudinale, transversale, prospective, rétrospective ?

Compte tenu des nuances que nous venons d'apporter précédemment, nous ne distinguerons pas les modalités d'enquêtes selon leur objectif, mais plutôt selon leur structure.

Les enquêtes longitudinales fournissent une information très riche, et surtout irremplaçable pour la mesure de l'incidence. Celle-ci est d'une grande valeur pour l'étude de la relation de cause à effet. Elles peuvent permettre des comparaisons du type "avant-après", qui bénéficient de la puissance statistique des dispositifs en séries appariées. Elles ont l'inconvénient de nécessiter la mise en oeuvre de moyens très lourds, puisque les observations sont répétées périodiquement. Du fait de la durée, les conditions d'observation initiales peuvent évoluer, affectant aussi bien les individus enquêtés, qui peuvent changer leurs pratiques, ou tout simplement sortir du champ d'observation (migration, mort, ou arrêt d'exploitation pour les élevages).

Les enquêtes transversales sont plus faciles à mettre en oeuvre, ce qui explique leur plus grande fréquence d'utilisation. Elles ne permettent pas la mesure de l'incidence, et ce n'est que dans des conditions particulières de relation parfaitement connue et stable entre prévalence et incidence que l'on peut en faire l'approximation.

Les enquêtes prospectives ont les avantages et les inconvénients des enquêtes longitudinales, telles que nous les avons présentées.

Les enquêtes rétrospectives sont fortement dépendantes de la qualité des informations selon le statut des individus. Il n'est pas exceptionnel que des sujets atteints, ou ayant été atteints mémorisent plus facilement une grande quantité d'informations, que les non-cas n'auront pas remarquées. Cette différence constitue une difficulté dans l'interprétation des écarts que l'on peut éventuellement constater.

Interprétation, évaluation des enquêtes

L'interprétation doit tout d'abord tenir compte de la manière dont le problème est posé, et dont la stratégie d'étude a été conçue. Elle doit ensuite établir la validité des faits observés (validité interne), avant d'envisager l'interprétation dans le sens des buts de l'enquête, et des finalités visées.

Définition du problème

Le but de l'étude doit être adapté aux besoins. Par exemple, avant de se préoccuper de la manière d'intervenir pour lutter contre une maladie, il faut tout d'abord se poser la question de la nécessité, et de l'opportunité de cette lutte. Une étude descriptive est par conséquent une priorité, avant d'autres types d'étude.

Les objectifs, qui correspondent à la formulation en termes quantifiables des buts visés, doivent être adaptés à ces buts. C'est la partie la plus délicate de la conception de l'enquête, car elle représente la concrétisation des choix stratégiques, ce qui justifie pleinement le temps qui peut être nécessaire à leur élaboration. Mieux vaut redéfinir les buts, si les objectifs ne peuvent être atteints.

La population cible doit être clairement définie au préalable, de façon à conditionner les modalités de sélection des échantillons, et partant, le champ d'extrapolation. Si l'on ne peut accéder qu'à des éleveurs inscrits sur telle liste, il ne sera pas possible d'extrapoler à l'ensemble des éleveurs de la région.

Si les buts concernent la distribution d'un phénomène dans le temps, ou dans l'espace, la formulation des objectifs et indicateurs épidémiologiques correspondants permet de concevoir le protocole adapté, et non l'inverse, à savoir, en fonction de données collectées pour un but donné, tenter, le plus souvent vainement, de produire une information qui va au delà des buts initiaux.

Les indicateurs épidémiologiques doivent correspondre aux objectifs. On peut être amené à redéfinir les objectifs, et éventuellement les buts, si l'on se rend compte que certains éléments constitutifs de ces indicateurs représentent une limite incontournable.

Par exemple, on ne saurait retenir des pourcentages d'animaux infectés comme pouvant représenter le pourcentage d'élevages infectés. Ou bien, la détection d'individus à sérologie positive ne correspond pas à la détection de la maladie.

Les étapes suivantes concernent la conception et l'exécution du protocole.

Validité interne

Les faits observés doivent refléter la situation réelle, et non résulter des conditions même d'observation, ce qui constitue alors un artefact. Chacun des points du protocole doit faire l'objet d'une évaluation, en vue de rechercher des biais potentiels.

Un biais est "une erreur systématique pouvant se produire à tout niveau d'une étude et dont la présence peut donner une fausse idée de la réalité" (11). Il modifie la valeur de la mesure, en affectant aussi bien une estimation (par exemple une prévalence), qu'une relation entre deux variables. Bon nombre de ces biais sont indépendants de l'échantillon, et par conséquent il faut distinguer les biais des fluctuations d'échantillonnage.

Il existe de très nombreux biais, que l'on peut regrouper en trois catégories : biais de sélection, biais d'information, et biais dus à un facteur de confusion.

Biais de sélection

La sélection des échantillons doit correspondre aux besoins du protocole : représentativité de la population cible pour une enquête descriptive, comparabilité des échantillons dans une étude analytique, ou simplement visant à une comparaison. La qualité de la sélection doit être évaluée tant à la phase de conception, que suite à l'exécution du protocole, qui peut ne pas être conforme aux prévisions initiales.

Une association peut être fallacieuse à cause d'un biais d'échantillonnage. Par exemple, pour étudier l'efficacité d'un traitement contre les mammites, on recrute des éleveurs sur la base de leur motivation à participer à l'enquête ; les éleveurs qui acceptent de se plier au protocole d'essai du nouveau traitement sont les plus motivés, mais aussi les plus attentifs dans le respect des techniques de santé ; si l'on n'y prend garde, et que les témoins sont choisis parmi une population différente par sa motivation, les différences éventuelles entre les deux lots ne pourront pas être rapportées qu'à l'effet du traitement étudié.

Biais d'information.

Les mesures doivent être réalisées avec des méthodes choisies en connaissance de leur qualités intrinsèques

(sensibilité, spécificité), et en vérifiant que leurs conditions de réalisation dans le cadre de cette étude ont permis d'atteindre le même degré de qualité. Dans le cas fréquent d'intervention de plusieurs enquêteurs, les procédures mises en oeuvre doivent avoir été standardisées, c'est-à-dire qu'elles sont reproduites selon les mêmes modalités. Pour cela, il faut les définir de façon explicite, et vérifier au préalable que les participants procèdent effectivement de la même façon, et obtiennent des résultats compatibles avec les marges d'erreur consenties. Une formation est le plus souvent nécessaire.

Un biais de mesure peut provenir de l'intervenant chargé de collecter l'information. L'erreur peut être liée à l'utilisation d'un instrument de mesure : par exemple, pour un cutimètre, appareil de mesure mal réglé (le cadran ne revient pas à 0 quand les mors sont jointifs), de marque et de performances différentes de celles des autres enquêteurs (cutimètre à pression du pouce versus à ressort isodynamométrique), ou encore utilisé dans des conditions inadéquates (mesure effectuée en tenant l'appareil verticalement versus horizontalement). Elle peut être d'ordre méthodologique et résulter d'une différence dans la définition de la mesure épidémiologique : par exemple, les cas cliniques à dénombrer ne sont pas définis par le même ensemble de symptômes selon les intervenants.

Ces biais ne sont détectables que si l'on a pris la précaution de récolter l'information correspondante nécessaire : identité de l'enquêteur connue pour toutes les observations, date, heure, ordre de collecte par exemple.

Biais dus à un facteur de confusion

Les biais de confusion, encore appelés biais de comparaison, modifient la relation entre le phénomène pathologique et le facteur de risque étudiés. Contrairement aux biais de sélection et de mesure, les biais dus à des facteurs de confusion peuvent être pris en compte, et corrigés quantitativement au moment de l'analyse statistique.

Ce facteur externe qui altère l'estimation de la relation entre facteur étudié et maladie est appelé facteur de confusion. C'est une "variable associée aussi bien à la variable indépendante d'intérêt particulier (facteur causal vrai) qu'à la variable dépendante (maladie)" (12). Il doit être neutralisé, contrôlé en estimant la relation entre le facteur étudié et la maladie pour chacun des niveaux d'exposition à ce facteur externe. On dit qu'un facteur externe est un facteur de confusion à partir du moment où les estimations du risque de maladie calculées à partir des données brutes et à partir des données corrigées en tenant compte du facteur externe sont différentes.

Par exemple, une étude (fictive) de type exposés/non-exposés montre que l'utilisation d'un tablier de caoutchouc par les vachers est associée à une plus faible incidence des mammites de la vache laitière, selon un risque

relatif brut de 4. L'examen du protocole ne révèle pas de faute dans la procédure d'échantillonnage. En revanche, une analyse complémentaire des données montre que le risque relatif associé au port du tablier, corrigé selon une variable externe portant sur la technicité des élevages, est voisin de 1 pour chacune des catégories distinguées selon le niveau de technicité. En revanche, le risque lié à ce facteur externe est de l'ordre de 6.

Un facteur de confusion peut augmenter, réduire, ou masquer la relation réelle entre le facteur d'exposition et la maladie.

Il peut aussi modifier la relation entre le facteur étudié et la maladie, dans le cas où il existe une interaction entre les deux facteurs, étudié et externe. On s'en rend compte dans le cas où les risques relatifs propres à chacune des catégories distinguées selon le niveau du facteur externe ne sont pas égaux. Pour reprendre notre exemple, ce serait le cas, si dans la population des élevages de meilleure technicité, le risque relatif était plus faible pour ceux utilisant le tablier, tandis qu'il resterait voisin de 1 pour les élevages de moindre technicité.

Cet exemple est caricatural, et chacun peut penser qu'il ne serait pas besoin de recourir à la statistique, et à la neutralisation des facteurs de confusion, pour se rendre compte d'évidences : n'importe qui penserait *a priori* au risque de liaison entre technicité et port de ce tablier, et risque de mammites. La facilité didactique de cet exemple permet de souligner, *a contrario*, la nécessité de ce type d'analyse statistique dans la grande majorité des circonstances où personne ne sait quel est au juste le facteur et l'effet, et même s'il y a une liaison *a priori*.

La prise en compte des facteurs de confusion peut être faite *a priori*, au moment de la planification de l'étude, ou au moment de l'analyse statistique.

Au moment de la planification, on procède au tirage au sort des groupes, dans le cas d'un essai. Dans les études d'observation, on peut recourir à la restriction de la population d'étude, ou à l'appariement des sujets par strate, ou individuellement.

On ne peut corriger un biais de confusion au moment de l'analyse statistique qu'en l'absence d'interaction. Dans le cas d'un seul facteur, on utilise la méthode d'ajustement de MANTEL-HAENSZEL. Dans les autres cas, on recourt à une modélisation qui permet la prise en compte de plusieurs facteurs de confusion. Les modèles les plus utilisés sont le modèle logistique, lorsque la maladie peut être codée en terme de présence/absence, et que l'on s'intéresse au risque de la maladie, et le modèle de COX, lorsqu'on suit des sujets au cours du temps et que l'on s'intéresse à l'incidence instantanée de la maladie (8). Dans les cas d'interaction, la relation globale entre le facteur étudié et la maladie n'a plus de sens, et l'on est conduit à poursuivre l'analyse et l'interprétation des résultats séparément pour chacune des catégories du facteur externe.

INTRODUCTION

Les résultats

Les résultats ne peuvent que correspondre aux objectifs, et surtout pas aller au delà, ce qui appartient alors au domaine de la spéculation. Ils sont prononcés pour l'échantillon étudié, et lui seul.

Extrapolation

Il faut tenir compte du fait que les résultats épidémiologiques, quels qu'ils soient, sont toujours "marqués" par leurs conditions d'obtention : ils ont été collectés à un moment donné, en un lieu donné, sur une population donnée.

Cette phase vise à extrapoler les résultats observés sur l'échantillon étudié à la population cible. Le pouvoir d'extrapolation dépend essentiellement de la qualité de représentativité des échantillons vis-à-vis des populations cibles : population cible pour une étude descriptive, problème de la relation causale étudiée en général pour une étude analytique, en fonction de la représentativité vis-à-vis des populations cibles exposées ou non.

Décisions pragmatiques

Etant donné qu'il n'existe pas d'étude épidémiologique parfaite (sinon on réaliserait des expériences), on doit prendre un certain nombre de décisions pour valider les résultats malgré les inconvénients constatés.

Action sur les biais

Les biais de sélection sont irrécupérables. En effet, s'ils résultent d'une attitude spontanée, c'est-à-dire d'un comportement, rien ne permet d'apprécier quels auraient été les comportements s'ils avaient été conformes aux attentes. On peut toutefois évaluer leur importance, soit par raisonnement, soit par enquête sur un échantillon complémentaire (par exemple des sujets n'ayant pas répondu ou ayant refusé de participer à l'enquête). De cette manière on peut apprécier si les résultats représentent une sous estimation, ou au contraire une surestimation, sans pour autant qu'il soit possible de quantifier les répercussions de ce biais.

On peut tenter de corriger les biais de mesure (à l'aide des valeurs de sensibilité et de spécificité connues) : on peut ainsi estimer la fréquence réelle à partir de la fréquence apparente.

Un biais de comparaison peut être compensé par des méthodes statistiques appropriées.

Point de vue pragmatique

La rigueur scientifique est indispensable pour éviter d'accorder une valeur illégitime à des informations corrompues par des conceptions, ou des modalités de récolte de mauvaise qualité. Mais les critères d'exigence absolue pour l'apport de connaissances nouvelles pourraient paralyser la mise en oeuvre d'études épidémiologiques

indispensables aux décisions de santé. C'est pourquoi, en définitive, il faut toujours revenir dans le processus d'interprétation aux buts visés, et aux finalités poursuivies, afin de relativiser l'importance des défauts constatés, en regard des décisions auxquelles ils sont sensés conduire.

CONCLUSION

L'épidémiologie se caractérise par sa méthode, au point qu'on lui reconnaît une manière particulière de penser.

Ses composantes méthodologiques lui permettent de fournir une information à valeur scientifique : leur maîtrise constitue aussi une excellente école de rigueur du raisonnement scientifique.

En plus, étant en prise directe avec les problèmes de santé qu'elle doit contribuer à résoudre, l'épidémiologie est, par essence même, une discipline tournée vers l'action. De ce fait, les résultats des enquêtes épidémiologiques doivent toujours être interprétés en fonction des décisions qu'elles doivent servir.

BIBLIOGRAPHIE

1. A.E.E.M.A. Qualité des analyses biologiques en épidémiologie et décisions de santé (Réunion A.E.E.M.A., Alfort, 7 Décembre 1989). *Epidémiol. Santé anim.*, 1990, 17. 144 p.
2. BERNARD (C.). Morceaux choisis et préfacés par J. Rostand. Paris, Gallimard, 1938. 285 p.
3. DAVENAS (E.) *et al.* Human basophil degranulation triggered by very dilute antisera against IgE. *Nature*, 1988, 333 : 816-818.
4. ELOIT (M.), KOUTCHOUKALI (M.A.). Sondage dans une population animale : estimation du taux d'infection des cheptels. *Epidémiol. Santé anim.*, 1984, 6 : 65-77.
5. FARVER (T.B.), THOMAS (C.), EDSON (R.K.). An application of sampling theory in animal disease prevalence survey design. *Prev. vet. Med.*, 1985, 3 : 463-473.
6. FARVER (T.B.). Disease prevalence estimation in animal populations using two-stage sampling designs. *Prev. vet. Med.*, 1987, 5 : 1-20.
7. JENICEK (M.) et CLEROUX (R.). *Epidémiologie. Principes, techniques applications.* Paris. Maloine, 1982. 454 p.
8. KLEINBAUM (D.G.), KUPPER (L.L.) MORGENSTERN (H.). *Epidemiologic research. Principles and quantitative methods.* New York, Van Nostrand Reinhold. 1982. 530 p.
9. MARTIN (S.W.), SHOUKRI (M.), THORBURN (M.). Evaluating the health status of herds based on tests applied to individuals. *Prev. vet. Med.*, 1992, 14 : 33-43.
10. ROTHMANN (K.J.). *Causal inference.* Chestnut Hill, Massachusetts, Epidemiology Resources Inc., 1988. 207 p.
11. SCHWARTZ (D.), FLAMANT (R.), LELLOUCH (J.). *L'essai thérapeutique chez l'Homme.* Paris, Flammarion, 1980. 288 p.
12. TOMA (B.), BENET (J.J.), DUFOUR (B.), ELOIT (M.), MOUTOU (F.), SANAA (M.). *Glossaire d'épidémiologie animale.* Maisons-Alfort, Editions du Point Vétérinaire, 1991. 365 p.

J.J. Bénet M. Sanaa B. Dufour B. Toma

BÉNET (J.J.), SANAA (M.), DUFOUR (B.), TOMA (B.). Methodology of surveys used in animal epidemiology. *Revue Elev. Méd. vét. Pays trop.*, 1993, **46** (3) : 403-422

In the present paper, the authors review the methodological bases of surveys, both descriptive and analytical, used in animal epidemiology. They give a comprehensive definition of the main terms applied to epidemiology as well as of the method used to evaluate the quality of these surveys.

Key words : Survey - Method - Epidemiology - Sampling - Evaluation.

BÉNET (J.J.), SANAA (M.), DUFOUR (B.), TOMA (B.). Metodología de las encuestas en epidemiología animal. *Revue Elev. Méd. vét. Pays trop.*, 1993, **46** (3) : 403-422

Los autores hacen una revision de las bases metodológicas de las encuestas, tanto descriptivas como analíticas, en la epidemiología animal. Se indican también las definiciones de los términos más utilizados en epidemiología, así como las modalidades de evaluación de la calidad de las encuestas.

Palabras claves : Encuesta - Metodo - Epidemiología - Muestreo - Evaluación.